(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 933 427 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
04.08.1999 Bulletin 1999/31

(51) Int. Cl.[6]: **C12N 15/53**, C12N 9/02, C12P 23/00

(21) Application number: **98309859.1**

(22) Date of filing: **02.12.1998**

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **02.12.1997 JP 33193697**

(71) Applicants:
- **Director General of National Institute of Fruit tree Science, Ministry of Agriculture, Forestry and Fisheries**
  Tsukuba-shi, Ibaraki 305-0852 (JP)
- **Bio-oriented Technology Research Advancement Institution**
  Omiya-shi, Saitama 331-0044 (JP)

(72) Inventors:
- **Yano, Masamitsu**
  Shimizu-shi, Shizuoka 424-0205 (JP)
- **Omura, Mitsuo**
  Shizuoka-shi, Shizuoka 422-8021 (JP)
- **Ikoma, Yoshinori**
  Shimizu-shi, Shizuoka 424-0902 (JP)
- **Komatsu, Akira**
  Shimizu-shi, Shizuoka 424-0204 (JP)

(74) Representative:
Woods, Geoffrey Corlett
J.A. KEMP & CO.
14 South Square
Gray's Inn
London WC1R 5LX (GB)

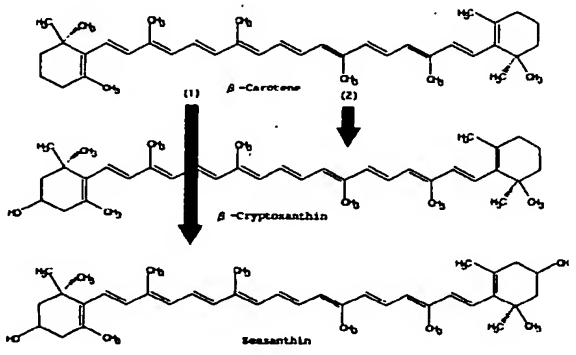
Remarks:
The applicant has subsequently filed a sequence listing and declared, that it includes no new matter.

(54) Beta-carotene hydroxylase gene

(57) A polypeptide which has β-carotene hydroxylase activity comprises an amino acid sequence selected from:

(i) the amino acid sequence of SEQ ID NO:2;
(ii) an amino acid sequence which is at least 70% homologous to the amino acid sequence of SEQ ID NO:2; and
(iii)a fragment of sequence (i) or (ii) which is not a said sequence (ii).

FIG. 1



EP 0 933 427 A2

Description

[0001] The present invention relates to a β-carotene hydroxylase, a DNA coding for the β-carotene hydroxylase, a recombinant vector comprising the DNA, a transformant transformed with the vector, a method for preparing the β-carotene hydroxylase and a method for preparing β-cryptoxanthin.

[0002] In carotenoids synthesized by animals, plants and microorganisms, there are a group of compounds with a hydroxyl group(s) generically called xanthophyll. These compounds are generated from carotenoids (starting substances) by the catalytic action of hydroxylase. For example, one hydroxyl group is introduced into β-carotene to yield β-cryptoxanthin, into which another hydroxyl group is introduced to yield zeaxanthin via the biosynthetic pathway shown below (see arrow (1) in Fig. 1):

β-carotene → β-cryptoxanthin → zeaxanthin

[0003] This β-cryptoxanthin is obtained by introducing a hydroxyl group into one of the two ionone rings present in β-carotene. When another hydroxyl group is introduced into a position symmetric to the former position, zeaxanthin is produced (Fig. 1).

[0004] In a large number of plants and microorganisms, metabolism proceeds from β-carotene to zeaxanthin, producing little β-cryptoxanthin, the intermediate into which only one hydroxyl group is introduced.

[0005] This reaction is controlled by a hydroxylase gene called Crt Z. In this enzyme reaction, it is considered that two hydroxyl groups are introduced almost simultaneously. For example, under the control of a hydroxylase gene cloned from a bacterium belonging to the genus Erwinia, zeaxanthin is produced which is obtainable by introducing two hydroxyl groups into β-carotene.

[0006] In Citrus unshiu (Satsuma mandarine) which is a major citrus fruit in Japan, β-cryptoxanthin obtainable by introducing one hydroxyl group into β-carotene is considered to be one of the most important carotenoids. In particular, β-cryptoxanthin occupies 60-70% of the total carotenoid content in the edible part of this fruit.

[0007] Considering this high β-cryptoxanthin content of Citrus unshiu, it is hard to think that the β-cryptoxanthin in Citrus unshiu is produced by a gene involved in the above-mentioned metabolic pathway. Also, it is still unknown whether β-cryptoxanthin is produced by those genes which have been already cloned.

[0008] It is an object of the present invention to provide a β-carotene hydroxylase and a gene coding for the enzyme.

[0009] As a result of intensive and extensive researches toward the solution of the above problem, the present inventors have succeeded in isolating from a citrus-derived cDNA library a DNA coding for a β-carotene hydroxylase. Thus, the present invention has been achieved.

[0010] Accordingly, the present invention provides a polypeptide which has β-carotene hydroxylase activity and which comprises an amino acid sequence selected from:

(i) the amino acid sequence of SEQ ID NO: 2;
(ii) an amino acid sequence which is at least 70% homologous to the amino acid sequence of SEQ ID NO: 2; and
(iii) a fragment of sequence (i) or (ii) which is not a said sequence (ii).

[0011] The polypeptide is suitably a protein, typically a recombinant protein, selected from:

(a) a protein consisting of the amino acid sequence of SEQ ID NO: 2; and
(b) a protein which consists of the amino acid sequence of SEQ ID NO: 2 having deletion, substitution or addition of one or several amino acids. The present invention further relates to:

- a polynucleotide, such as a DNA, coding for a polypeptide of the invention;
- a DNA coding for a β-carotene hydroxylase, comprising the nucleotide sequence of SEQ ID NO: 1 and preferably the coding portion of that sequence;
- a recombinant vector comprising a polynucleotide of the invention, such as an expression vector in which the polynucleotide is operably linked to a promoter;
- a transformant which is transformed with the above vector, suitably a procaryote or a eucaryote that is not a human-being;
- a method for preparing a β-carotene hydroxylase and/or β-cryptoxanthin, which method comprises maintaining a transformant as defined above under conditions such that the desired β-carotene hydroxylase and/or β-cryptoxanthin is expressed and recovering the β-carotene hydroxylase and/or β-cryptoxanthin;
- a method for preparing a β-carotone hydroxylase comprising culturing the above transformant in a medium and recovering the β-carotene hydroxylase from the resultant culture; and
- a method for preparing β-cryptoxanthin comprising culturing the above transformant in a medium and recover-

ing β-cryptoxanthin from the resultant culture.

[0012] In the accompanying drawings:

Fig. 1 is a diagram showing a biosynthetic pathway of carotenoids.
Fig. 2 presents chromatograms showing the results of high performance liquid chromatography.
Fig. 3 is a diagram showing comparison of homology in amino acid sequences between the β-carotene hydroxylase of the invention and other enzymes.

[0013] A polypeptide of the invention consists essentially of the amino acid sequence set out in SEQ ID NO: 2 or a substantially homologous sequence, or of a fragment of either of these sequences. The polypeptide may consist of such a sequence. In general, the naturally occurring sequence shown in SEQ ID NO: 2 is preferred. However, the polypeptides of the invention include homologues of the natural sequence, and fragments of the natural sequence and of its homologues, which are capable of functioning as a β-carotene hydroxylase. The β-carotene hydroxylase catalyzes the reaction indicated by arrow (2) in Fig. 1. It typically does not catalyze the reaction indicated by arrow (1).

[0014] An amino acid sequence at least 70% homologous to the amino acid sequence of SEQ ID NO: 2 will be preferably at least 80 or 90% and more preferably at least 95, 97 or 99% homologous thereto. Homology is calculated on the basis of amino acid identity. SEQ ID NO: 2 can thus be modified, typically by deletion, substitution and/or addition. One or several, for example up to five, amino acids of SEQ ID NO: 2 may be modified in this way. In fact more may be modified in this way within the scope of the invention.

[0015] Amino acid substitutions may be made, for example from 1, 2 or 3 to 10, 20 or 30 substitutions. The modified polypeptide retains β-carotene hydroxylase activity. Conservative substitutions may be made, for example according to the following Table. Amino acids in the same block in the second column and preferably in the same line in the third column may be substituted for each other:

ALIPHATIC	Non-polar	G A P
		I L V
	Polar - uncharged	C S T M
		N Q
	Polar - charged	D E
		K R
AROMATIC		H F W Y

[0016] One or more amino acid residues of SEQ ID NO: 2 may alternatively or additionally be deleted. From 1, 2 or 3 to 10, 20 or 30 residues may be deleted, or more. Polypeptides of the invention also include fragments of the above-mentioned sequences (i) and (ii). Such fragments retain β-carotene hydroxylase activity. Fragments may be at least from 10, 12, 15 or 20 to 60, 100 or 200 amino acids in length.

[0017] One or more amino acids may be alternatively or additionally added to SEQ ID NO: 2. An extension may be provided at the N-terminus or C-terminus of the sequence of SEQ ID NO: 2. The or each extension may be quite short, for example from 1 to 10 amino acids in length. Alternatively, the extension may be longer. A carrier protein may be fused to an amino acid sequence according to the invention. A fusion protein incorporating the sequence (i), (ii) or (iii) above can thus be provided.

[0018] A polypeptide of the invention may thus have been modified for example by the addition of Histidine residues or a T7 tag to assist in identification or purification of the polypeptide. A signal sequence may have been added to promote secretion of the polypeptide from a cell.

[0019] Polynucleotides of the invention may be single-stranded or double-stranded. They may be a DNA such as a cDNA, or an RNA. They thus consist essentially of DNA or RNA encoding a polypeptide of the invention such as the protein having the amino acid sequence of SEQ ID NO: 2. A DNA of the invention can consist essentially of the coding sequence of SEQ ID NO: 1 or of a modified form of that sequence, in particular a modified form that is substantially homologous to the coding sequence of SEQ ID NO: 1.

[0020] A DNA sequence of the invention may therefore be at least 80 or 90%, and more preferably at least 95%, homologous to the coding sequence of SEQ ID NO: 1. There may be as few as from 1 to 30, for example from 5 to 20,

nucleotide differences. A DNA sequence may code for the amino acid sequence of SEQ ID NO: 2 but one or more codons may be different from corresponding codons of the coding sequence of SEQ ID NO: 1 due the degeneracy of the genetic code.

[0021] A polynucleotide of the invention is typically capable of hybridising selectively with the sequence shown in SEQ ID NO: 1 which is complementary to the coding sequence of SEQ ID NO: 1. A polynucleotide of the invention and the complementary sequence of SEQ ID NO: 1 can hybridize at a level significantly above background. Background hybridization may occur, for example, because of other cDNAs present in a cDNA library. The signal level generated by the interaction between a polynucleotide of the invention and the sequence shown in SEQ ID NO: 1 which is complementary to the coding sequence is typically at least 10 fold, preferably at least 100 fold, as intense as interactions between other polynucleotides and that complementary sequence. The intensity of interaction may be measured, for example, by radiolabelling the probe, e.g. with ^{32}P. Selective hybridisation is typically achieved using conditions of medium to high stringency (for example 0.03M sodium chloride and 0.03M sodium citrate at from about 50°C to about 60°C).

[0022] Polynucleotides of the invention may be polynucleotides which include within them synthetic or modified nucleotides. A number of different types of modification to polynucleotides are known in the art. These include methylphosphonate and phosphorothioate backbones, addition of acridine or polylysine chains at the 3' and/or 5' ends of the molecule. For the purposes of the present invention, it is to be understood that the polynucleotides described herein may be modified by any method available in the art. Such modifications may be carried out in order to enhance the *in vivo* activity or lifespan of polynucleotides of the invention.

[0023] Polynucleotides according to the invention may be produced recombinantly, synthetically, or by any means available to those of skill in the art. They may also be cloned by standard techniques. The polynucleotides are typically provided in isolated and/or purified form.

[0024] A cDNA of the invention can be isolated by the following procedures. Briefly, a primer is designed based on a conserved region of a gene coding for a bacterium-derived β-carotene hyroxylase. Then, 3' RACE RT-PCR is performed using the above primer and, as a template, a first strand cDNA from the fruit (juice sacs) and flower of Citrus unshiu (variety: Miyagawa early) to obtain a cDNA fragment of the Citrus unshiu β-carotene hydroxylase. Subsequently, using this cDNA fragment as a probe, the β-carotene hydroxylase of interest can be isolated from a cDNA library derived from the edible part of Citrus unshiu.

1. Cloning of a DNA Coding for the β-Carotene Hydroxylase (1) Preparation of primers

[0025] First, primers for use in the 3' RACE RT-PCR to be described later are prepared. In order to design a primer that is more specific to a DNA of interest, it is appropriate to prepare an oligonucleotide coding for a region in which amino acid residues are highly conserved among various bacteria and plants. Such a primer can be prepared by conventional chemical synthesis. For example, the following amino acid sequences may be selected as regions satisfying the above-mentioned condition:

i) ⌈Phe Glu Leu Asn Asp Val Phe Ala⌋ (SEQ ID NO: 3)
ii) ⌈His Asp Gly Leu Val His⌋ (SEQ ID NO: 4)

[0026] Since these two regions with highly conserved amino acid residues are located close to each other, they cannot be used as a sense primer and an antisense primer in a PCR. Thus, in the present invention, 3' RACE RT-PCR method was employed in which each of these sequences was used as a sense primer.

[0027] The above sequences are found within the amino acid sequences for an Arabidopsis-derived and an Erwinia-derived β-carotene hydroxylase described by Zairen Sun et al., The Journal of Biological Chemistry, 1996; Vol. 271, No. 40; 24349-24352 and Nakagawa M. and N. Misawa, Agric. Biol. Chem. 55:2147-2148, respectively.

[0028] Based on these amino acid sequences, oligonucleotide primers with the following sequences, for example, are prepared. However, the primers are not limited to these sequences.

Sense 1 primer (Bech-a): TT(t/c)GA(g/a)CTAAA(c/t)GA(t/c)GTN (SEQ ID NO: 5)
Sense 2 primer (Bech-B): CACGA(c/t)GGTCTNGTNCA (SEQ ID NO: 6)

(2) 3' RACE RT-PCR

[0029] Subsequently, a 3'RACE RT-PCR is performed using the two sense primers synthesized. RT-PCR (reverse transcription-PCR) is a method in which a DNA is synthesized (reverse transcribed) with RNA as a template using a reverse transcriptase, and thereafter a PCR is performed using the synthesized DNA as a template. 3' RACE (rapid amplification of cDNA ends) is a method in which an RT-PCR is performed based on a nucleotide sequence of a known region to thereby clone the unknown region of a cDNA of interest up to the relevant cDNA end.

[0030] First, a reverse transcription is performed using an oligo(dT) primer having an adaptor sequence at its 5' end to thereby synthesize a first strand cDNA. All of the resultant first strand cDNA molecules have a structure in which the adaptor sequence is attached to the end. Therefore, in the cDNA to be cloned, the unknown region is located between the known sequence and the adaptor sequence. Then, the unknown region (cDNA partial sequence) sandwiched between the two sequences can be amplified by performing a PCR using a part of the known sequence as a sense primer together with the adaptor primer.

[0031] An RT-PCR can be performed using a commercial kit (T-Primed First-Strand Kit: Pharmacia).

(3) Preparation of a cDNA Library

[0032] In order to obtain the full-length cDNA of interest from a fruit-derived cDNA library using the cDNA partial sequence obtained above as a probe, the library is prepared as described below.

[0033] Total RNA is isolated from individual citrus organs or tissues (fruit, leaf, root, flower, callus, etc.) using a guanidine reagent or SDS-phenol. Then, mRNA is prepared from the total RNA by the affinity column method using oligo dT-cellulose or poly U-Sepharose carried on Sepharose 2B or by a method using an oligotex resin. Using the resultant mRNA as a template, a single-stranded cDNA is synthesized with a reverse transcriptase. Thereafter, a double-stranded cDNA is synthesized from the single-stranded cDNA. The resultant double-stranded cDNA is ligated to an appropriate plasmid or phage vector using a ligase to thereby obtain a recombinant DNA. By infecting or transforming Escherichia coli or the like with this recombinant DNA, a cDNA library capable of screening by plaque or colony hybridization can be obtained.

(4) Isolation of a β-Carotene Hydroxylase cDNA Homologue from the cDNA Library

[0034] Subsequently, screening for the full-length cDNA sequence is performed by plaque or colony hybridization using the cDNA sequence isolated by the 3' RACE RT-PCR described above as a probe. For this hybridization, a commercial kit such as ECL Nucleic Acid Labelling and Detection System (Amersham) may be used.

(5) Determination of the Nucleotide Sequence

[0035] The nucleotide sequence of the obtained clone is determined. This can be performed by conventional methods such as Maxam-Gilbert method, the dideoxy method or the like. Usually, the determination is carried out with an automatic DNA sequencer.

[0036] SEQ ID NO: 1 shows the nucleotide sequence for the DNA of the invention and SEQ ID NO: 2 shows the amino acid sequence for the β-carotene hydroxylase of the invention. However, as long as a protein consisting of this amino acid sequence has β-carotene hydroxylase activity, the sequence may have some mutation such as deletion, substitution or addition of one or several amino acids. For example, a protein consisting of the amino acid sequence of SEQ ID NO: 2 in which Met at the first position has been deleted is also included in the protein of the invention.

[0037] The β-carotene hydroxylase activity in the present invention means an activity to perform a catalytic reaction producing β-cryptoxanthin from β-carotene.

[0038] Once the nucleotide sequence for the DNA of the invention has been established, the DNA of the invention can be obtained by chemical synthesis or by hybridization using a DNA fragment having a part of the sequence as a probe.

2. Preparation of a Recombinant Vector and a Transformant

(1) Preparation of a Recombinant Vector

[0039] The recombinant vector of the invention can be obtained by ligating (inserting) the DNA of the invention to (into) an appropriate vector. The vector into which the DNA of the invention is to be inserted is not particularly limited as long as it is replicable in a host. For example, a plasmid DNA, a phage DNA or the like may be used.

[0040] A plasmid DNA can be prepared from E. coli or Agrobacterium by alkali extraction (Birnboim, H.C. & Doly, J. (1979), Nucleic Acid Res., 7:1513) or variations thereof. Alternatively, a commercial plasmid such as pBluescript II SK+ (Stratagene), pUC118 (TaKaRa), pUC119 (TaKaRa), pGEM-T (Promega) or the like may be used. It is preferred that these plasmids contain a selectable marker such as ampicillin resistance gene, kanamycin resistance gene or chloramphenicol resistance gene.

[0041] As a phage DNA, M13mp18, M13mp19 or the like may be given.

[0042] For insertion of the DNA of the invention into a vector, a method may be employed in which the purified DNA is digested with an appropriate restriction enzyme and then inserted into the relevant restriction site or the multi-cloning

site of the vector for ligation. The DNA of the invention should be incorporated in the vector in such a manner that the function thereof is operable. For this purpose, the vector of the invention may contain a terminator, ribosome binding sequence or the like in addition to a promoter and the DNA of the invention.

(2) Preparation of a Transformant

[0043] The transformant of the invention can be obtained by introducing the recombinant vector of the invention into a host so that the gene of interest can be expressed.

[0044] The host is not particularly limited as long as it can express the DNA of the invention. Specific examples of the host include Escherichia or Bacillus bacteria such as E. coli and Bacillus subtilis; yeasts such as Saccharomyces cerevisiae; or animal cells such as COS cells and CHO cells.

[0045] When a bacterium such as E. coli is used as the host, preferably, the recombinant vector of the invention is capable of autonomous replication in the host and, at the same time, is constituted by a promoter, a ribosome binding sequence, the DNA of the invention and a transcription termination sequence. The vector may also contain a gene to control the promoter.

[0046] As the expression vector, pBluescript II vector (Stratagene), pET vector (Stratagene) or the like may be used.

[0047] As the promoter, any promoter may be use as long as it can direct the expression of the DNA of the invention in the host such as E. coli.

[0048] For example, an E. coli-derived or phage-derived promoter such as trp promoter, lac promoter, P_L promoter or P_R promoter may be used.

[0049] As a method for introducing the recombinant vector into the bacterium, any method of DNA introduction into bacteria may be used. For example, a method using calcium ions (Proc. Natl. Acad. Sci., USA, 69:2110-2114 (1972)) may be used. When a yeast is used as the host, YEp13, YEp24, YCp50 or the like is used as an expression vector. As a promoter used in this case, any promoter may be used as long as it can direct the expression of the DNA of the invention in yeasts. For example, gal1 promoter, ga110 promoter, heat shock protein promoter, MF α1 promoter or the like may be enumerated.

[0050] As a method for introducing the recombinant vector into the yeast, any method of DNA introduction into yeasts may be used. For example, electroporation (Methods Enzymol., 194:182-187 (1990)), the spheroplast method (Proc. Natl. Acad. Sci., USA, 84:1929-1933 (1978)), the lithium acetate method (J. Bacteriol., 153:163-168 (1983)) or the like may be enumerated.

[0051] When an animal cell is used as the host, pcDNAI/Amp (Invitrogen) or the like is used as an expression vector. In this case, the early gene promoter of human cytomegalovirus or the like may be used as a promoter.

[0052] As a method for introducing the recombinant vector into the animal cell, electroporation, the calcium phosphate method, lipofection or the like may be enumerated.

[0053] The recombinant vector of the invention incorporated in E. Coli (designation: EpCitBECH1) was deposited on 1 December 1997 at the National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology (1-3, Higashi 1-Chome, Tsukuba City, Ibaraki Pref., Japan) as FERM BP-6188 under the Budapest Treaty. The deposit was made in the name of Toru Maotani; Director General of National Institute of Fruit Tree Science, Ministry of Agriculture, Forestry and Fisheries; 2-1, Fujimoto; Tsukuba-shi; Ibaraki 305; Japan.

3. Production of the β-Carotene Hydroxylase

[0054] The β-carotene hydroxylase of the invention can be obtained by culturing the transformant described above and recovering the β-carotene hydroxylase from the resultant culture.

[0055] The cultivation of the transformant of the invention in a medium is carried out by conventional methods used for culturing a host.

[0056] As a medium to culture the transformant obtained from a microorganism host such as E. coli or yeast, either a natural or a synthetic medium may be used as long as it contains carbon sources, nitrogen sources and inorganic salt sources assimilable by the microorganism and can be used for effective cultivation of the transformant.

[0057] As carbon sources, carbohydrates such as glucose, fructose, sucrose, starch; organic acids such as acetic acid, propionic acid; and alcohols such as ethanol and propanol may be used.

[0058] As nitrogen sources, ammonia; ammonium salts of inorganic or organic acids such as ammonium chloride, ammonium sulfate, ammonium acetate, ammonium phosphate; other nitrogen-containing compounds; Peptone; meat extract; corn steep liquor and the like may be used.

[0059] As inorganic substances, potassium dihydrogen phosphate, dipotassium hydrogen phosphate, magnesium phosphate, magnesium sulfate, sodium chloride, iron(II) sulfate, manganese sulfate, copper sulfate, calcium carbonate and the like may be used.

[0060] Usually, the cultivation is carried out under aerobic conditions (such as shaking culture or aeration agitation

culture) at 28 °C for 48 to 60 hrs. During the cultivation, the pH is maintained at 7.0 to 7.5. The pH adjustment is carried out using an inorganic or organic salt, an alkali solution or the like. When an E. coli transformant is cultured, it is preferable to allow pACCAR16ΔcrtX plasmid (having 4 Erwinia-derived genes that can produce carotenoids from farnesyl diphosphate to β-carotene) to coexist in the E. coli.

[0061] During the cultivation, an antibiotic such as ampicillin or tetracycline may be added to the medium if necessary.

[0062] When a microorganism transformed with an expression vector using an inducible promoter is cultured, an inducer may be added to the medium if necessary. For example, when a microorganism transformed with an expression vector using Lac promoter is cultured, isopropyl-β-D-thiogalactopyranoside (IPTG) or the like may be added. When a microorganism transformed with an expression vector using trp promoter is cultured, indoleacetic acid (IAA) or the like may be added.

[0063] As a medium to culture a transformant obtained from an animal cell as a host, commonly used RPMI1640 medium or DMEM medium, or one of these media supplemented with fetal bovine serum, etc. may be used.

[0064] Usually, the cultivation is carried out in the presence 5% CO_2 at 37 °C for 1 to 2 days.

[0065] During the cultivation, an antibiotic such as kanamycin or penicillin may be added to the medium if necessary.

[0066] After the cultivation, the β-carotene hydroxylase of the invention is recovered by disrupting the microorganisms or cells if the enzyme is produced in the microorganisms or cells. If the β-carotene hydroxylase of the invention is produced outside of the microorganisms or cells, the culture fluid (as it is or after centrifugation to remove the microorganisms or cells) is subjected to conventional biochemical techniques used for isolating/purifyinng a protein. These techniques include ammonium sulfate precipitation, gel chromatography, ion exchange chromatography and affinity chromatography. These techniques may be used independently or in an appropriate combination to isolate and purify the β-carotene hydroxylase of the invention from the culture.

[0067] The confirmation that the finally obtained protein is a β-carotene hydroxylase can be made by SDS-polyacrylamide gel electrophoresis.

4. Production of β-Cryptoxanthin

[0068] In the present invention, it is also possible to produce β-cryptoxanthin in the same manner as described in the purification of the β-carotene hydroxylase. Briefly, the transformant described above is cultured in a medium and then β-cryptoxanthin is extracted from the resultant culture. The method of cultivation is the same as described in "3. Production of the β-Carotene Hydroxylase".

[0069] After the cultivation, the microorganisms or cells are removed from the culture by centrifugation of the like. Then, β-cryptoxanthin can be extracted from the culture by HPLC or the like.

[0070] The confirmation that the finally extracted substance is β-cryptoxanthin can be made by ^{1}H-NMR, ultraviolet-visible spectroscopy, mass spectrometry, etc.

[0071] Now, the present invention will be described more specifically below with reference to the following Examples, which should not be instrued as limiting the technical scope of the invention.

EXAMPLE 1

Cloning of a cDNA Coding for the β-Carotene Hydroxylase

(1) Cloning of a Partial cDNA of Interest Using 3' RACE RT-PCR

[0072] A 1st-strand cDNA was prepared by performing a reverse transcription using NotI-D$(T)_{18}$ (5'd[AACTGGAAGAATTCGCGGCCGCAGGAAT$_{18}$]3')(SEQ ID NO: 7) as a primer and RNA from the fruit (juice sacs) and flower of Citrus unshiu (variety: Miyagawa early) as a template. At the time of this synthesis, NotI adaptor sequence (TGGAAGAATTCGCGGCCGCAG) (SEQ ID NO: 8) was added at the 3' end of every 1st-strand cDNA fragment. Using this 1st-strand cDNA as a template, a PCR was performed with Sense 1 primer and the adaptor primer. The reaction was carried out 35 cycles, 1 cycle consisting of denaturation at 94.5°C for 40 sec and annealing/extension at 60 °C for 2 min. In this first stage PCR, however, the adaptor primer used is the sequence contained commonly in all of the cDNA fragments generated by the reverse transcription reaction. Therefore, the PCR product obtained at this stage contains a large number of non-specifically amplified DNA fragments. In order to amplify the DNA of interest specifically, the second stage PCR was performed using Sense 2 primer. This reaction was carried out 35 cycles, one cycle consisting of denaturation at 94.5°C for 35 sec, annealing at 55°C for 45 sec and extension at 72 °C for 1 min. For the RT-PCR, a commercial kit (T-Primed First-Strand Kit: Pharmacia) was used.

[0073] By the above procedures, a cDNA partial sequence coding for a citrus β-carotene hydroxylase was obtained.

(2) Preparation of a cDNA Library from a Citrus Fruit Tissue

[0074] Total RNA was isolated from the fruit (juice sac tissue) of Citrus unshiu (variety: Miyagawa early) using guanidine thiocyanate. After the isolated total RNA was purified into mRNA using Oligotex-dT30 [Super] (TaKaRa), a first-strand cDNA was synthesized using an oligo$(dT)_{12-18}$ primer and a reverse transcriptase from Moloney murine leukemia virus (MMLV). Further, a second-strand cDNA was synthesized using a DNA polymerase (Pharmacia). To the resultant double-stranded cDNA, EcoRI adaptor was added by T4 DNA ligase, followed by ligation to Uni-ZAP EcoRI phagemid vector (Stratagene).

(3) Screening of the Full-Length cDNA of Interest by Plaque Hybridization

[0075] Subsequently, the full-length cDNA sequence of interest was screened by plaque hybridization using the cDNA partial sequence coding for a citrus β-carotene hydroxylase obtained by the 3' RACE RT-PCR described above.

[0076] A commercial kit (ECL nucleic acid labelling and detection system: Amersham) was used for the hybridization. As a result of the screening ($3x10^4$ pfc), a β-carotene hydroxylase cDNA homologue of 1158 bp in full length was isolated which was coding for a peptide of 311 amino acid residues with an estimated molecular weight of 34.7 kDa. This clone exhibited 76.3% homology to an Arabidopsis-derived β-carotene hydroxylase cDNA which produces zeaxanthin from β-carotene, and 35.7 to 39.8% homology to bacteria-derived β-carotene hydroxylase genes which produce zeaxanthin. This clone was designated "CitBECH1". The nucleotide sequence for CitBECH1 is shown in SEQ ID NO: 1, and the amino acid sequence encoded by CitBECH1 is shown in SEQ ID NO: 2.

[0077] The results of comparison of homology between conventional β-carotene hydroxylases and the β-carotene hydroxylase of the invention are shown in Fig. 3.

[0078] In Fig. 3, shown at the top row (CitBECH1) is the β-carotene hydroxylase amino acid sequence encoded by the gene of the invention. The others are amino acid sequences encoded by related genes; any of these sequences is a sequence for a gene producing zeaxanthin from β-carotene skipping over β-cryptoxanthin.

EXAMPLE 2

Production of β-Cryptoxanthin in Escherichia coli Having the β-Carotene Hydroxylase

(1) Expression of the DNA of the Present Invention

[0079] The isolated clone was inserted into pBluescript II SK+ plasmid having an ampicillin resistance gene. The resultant plasmid was introduced into Escherichia coli, in which pACCAR16ΔcrtX plasmid (having 4 Erwinia-derived genes that can produce from farnesyl diphosphate to β-carotene) was allowed to coexist. The resultant E. coli was cultured in LB medium at 28°C for 60 hrs.

[0080] Then, the culture was subjected to acetone extraction. The acetone extract from the transformant was subjected to HPLC using a system manufactured by Japan Spectroscopic Co., Ltd. As a column, a C30 column manufactured by YMC was used. As eluent A, a mixture of methanol/methyl-t-butyl ether/water mixed at a ratio of 81/15/4 was used. As eluent B, a mixture of methanol/methyl-t-butyl ether mixed at a ratio of 10/90 was used. Gradient conditions were as follows: eluent A 100% at the time of start; eluent A 20% and eluent B 80% 70 min after the start. The flow rate was 1.0 ml/min and the column temperature 22°C. The detection wave length was 450 nm.

[0081] As a result, the chromatograms shown in Fig. 2 were obtained. When the resultant peaks were compared with the peaks of the carotenoid standard products manufactured by Funakoshi, it was found that the E. coli produced β-cryptoxanthin, β-carotene and zeaxanthin at a ratio of 43:22:11. From this result, it was judged that the citrus-derived β-carotene hydroxylase mainly produces β-cryptoxanthin.

(2) Production and Identification of β-Cryptoxanthin

[0082] Plasmid pCitBECH 1-introduced, β-carotene-producing E. coli JM101 [E. coli (PACCAR16ΔcrtX, pCitBECH 1)] (presenting a yellow color) was cultured in 1.6 L of 2xYT medium [1.6% tryptone, 1% yeast extract, 0.5% NaCl] containing 150 μg/ml of ampicillin (Ap) and 30 μg/ml of chloramphenicol (Cm) at 30°C for 28 hrs. Cells were harvested from the culture fluid were subjected to extraction with 360 ml of acetone. The resultant extract was concentrated and extracted with 200 ml of chloroform/methanol (9/1) twice, followed by concentration and drying. The resultant solid material was dissolved in a small amount of chloroform/methanol (9/1) and then subjected to thin layer chromatography (TLC) in which the sample was developed with chloroform/methanol on a silica gel preparative TLC plate from Merck.

[0083] As a result of this TLC, the initial pigments were divided into two spots of Rf values 0.4 (dark) and 0.1 (very light), respectively, in addition to the β-carotene spot at the top. Then, the dark yellow pigment of Rf 0.4 was scratched

off from the TLC plate, dissolved in a small amount of chloroform/methanol (1/1) and subjected to TOYOPEARL HW-40 column chromatography for development and elution.

[0084] As a result, 1 mg of the pure pigment was obtained.

[0085] This pigment was considered to be β-cryptoxanthin from the results of examination of the ultraviolet-visible spectrum (λ 425, 448, 475 nm in methanol) and the FD-MS spectrum (m/e 553, [M]+). Further, two signals of 3-hydroxy-β-ionone ring and β-ionone ring (G. Englert, N.M.R. of Carotenoids edited by G. Britton, T.W. Goodwin, Carotenoid Chemistry and Biochemistry) were confirmed from its ^{1}H-NMR spectrum.

[0086] Consequently, this pigment was identified as β-cryptoxanthin (Fig. 2). Fig. 2 indicates that the gene of the invention is involved in biosynthesis of β-cryptoxanthin. In Fig. 2, the top panel shows the results of HPLC analysis of the carotenoids produced by E. coli in which an Erwinia-derived β-carotene biosynthesis gene was incorporated; the middle panel shows the results of HPLC analysis of the carotenoids produced by the above E. coli in which the gene of the invention was further incorporated; and the bottom panel shows the results of HPLC analysis of zeaxanthin, β-cryptoxanthin and β-carotene standard products.

[0087] From Fig. 2, it can be seen that, different from conventional β-carotene hydroxylases encoded by known genes (Crt Z) derived from Erwinia and marine bacteria, the β-carotene hydroxylase encoded by the gene of the invention catalyzes synthesis of carotenoids in such a manner that β-cryptoxanthin is produced mainly and zeaxanthin is produced in only a small amount (Fig. 2, middle panel).

[0088] According to the present invention, a β-carotene hydroxylase, a DNA coding for the β-carotene hydroxylase, a recombinant vector comprising the DNA, a transformant transformed with the vector, a method for preparing the β-carotene hydroxylase and a method for preparing β-cryptoxanthin are provided.

[0089] The β-carotene hydroxylase of the invention is useful in catalyzing synthesis of β-cryptoxanthin, a pigment necessary and important for maintaining the quality and function of citrus fruits and processed products thereof.

SEQUENCE LISTING

SEQ ID NO: 1

SEQUENCE LENGTH: 1158

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

ORIGINAL SOURCE

ORGANISM: Citrus unshiu

FEATURE

NAME/KEY: CDS

LOCATION: 87..1019

SEQUENCE DESCRIPTION

```
CCACAATCCA CTTCACATCA ACTCTTCCTC TTTTCAAGTG CTTTTACTCT AAAACCCAAA   60
ACCTCGTAAA CAAACAAAAC CCCACC ATG GCG GTC GGA CTA TTG GCC GCC ATA   113
                              Met Ala Val Gly Leu Leu Ala Ala Ile
                                1               5
GTC CCG AAG CCC TTC TGT CTC CTC ACA ACA AAA CTT CAA CCC TCT TCG    161
Val Pro Lys Pro Phe Cys Leu Leu Thr Thr Lys Leu Gln Pro Ser Ser
10                  15                  20                  25
CTC CTC ACA ACA AAA CCC GCT CCC CTT TTT GCC CCT CTC GGT ACC CAC    209
Leu Leu Thr Thr Lys Pro Ala Pro Leu Phe Ala Pro Leu Gly Thr His
                30                  35                  40
CAT GGC TTC TTT AAT GGC AAA AAC CGA AGA AAA CTC AAC TCT TTC ACC    257
His Gly Phe Phe Asn Gly Lys Asn Arg Arg Lys Leu Asn Ser Phe Thr
            45                  50                  55
GTA TGT TTT GTT TTA GAG GAG AAA AAA CAA AGC ACC CAG ATC GAG ACT    305
Val Cys Phe Val Leu Glu Glu Lys Lys Gln Ser Thr Gln Ile Glu Thr
```

```
                60                  65                  70
TTC ACG GAC GAG GAG GAG GAG GAG TCG GGT ACC CAG ATC TCG ACT GCT    353
Phe Thr Asp Glu Glu Glu Glu Glu Ser Gly Thr Gln Ile Ser Thr Ala
     75                  80                  85
GCC CGC GTG GCC GAG AAA TTG GCG AGA AAG AGA TCC GAG AGG TTC ACT    401
Ala Arg Val Ala Glu Lys Leu Ala Arg Lys Arg Ser Glu Arg Phe Thr
90                  95                 100                 105
TAT CTC GTT GCT GCC GTC ATG TCT AGT TTT GGT ATC ACT TCC ATG GCT    449
Tyr Leu Val Ala Ala Val Met Ser Ser Phe Gly Ile Thr Ser Met Ala
                110                 115                 120
GTC ATG GCT GTT TAT TAC AGG TTC TGG TGG CAA ATG GAG GGT GGA GAG    497
Val Met Ala Val Tyr Tyr Arg Phe Trp Trp Gln Met Glu Gly Gly Glu
            125                 130                 135
GTG CCT TTA GCT GAA ATG TTT GGC ACA TTT GCT CTC TCT GTT GGT GCT    545
Val Pro Leu Ala Glu Met Phe Gly Thr Phe Ala Leu Ser Val Gly Ala
        140                 145                 150
GCT GTG GGC ATG GAG TTT TGG GCA CGA TGG GCT CAT AAA GCT CTG TGG    593
Ala Val Gly Met Glu Phe Trp Ala Arg Trp Ala His Lys Ala Leu Trp
    155                 160                 165
CAT GCT TCT TTA TGG CAT ATG CAC GAG TCT CAC CAT CGA CCA AGA GAG    641
His Ala Ser Leu Trp His Met His Glu Ser His His Arg Pro Arg Glu
170                 175                 180                 185
GGT CCT TTT GAG CTA AAC GAT GTG TTT GCC ATA ATC AAC GCA GTT CCA    689
Gly Pro Phe Glu Leu Asn Asp Val Phe Ala Ile Ile Asn Ala Val Pro
                190                 195                 200
GCC ATA GCC CTT CTC TCT TTT GGC TTC TTC CAC AAA GGC CTT GTA CCT    737
Ala Ile Ala Leu Leu Ser Phe Gly Phe Phe His Lys Gly Leu Val Pro
            205                 210                 215
GGT CTC TGC TTT GGT GCT GGA CTT GGC ATT ACG GTG TTT GGG ATG GCC    785
```

```
Gly Leu Cys Phe Gly Ala Gly Leu Gly Ile Thr Val Phe Gly Met Ala
        220                 225                 230

TAC ATG TTC GTC CAC GAT GGT CTC GTT CAC AAA AGG TTC CCT GTG GGT   833
Tyr Met Phe Val His Asp Gly Leu Val His Lys Arg Phe Pro Val Gly
    235                 240                 245

CCC ATT GCC GAC GTG CCT TAT TTC CGG AGA GTC GCT GCG GCT CAC CAG   881
Pro Ile Ala Asp Val Pro Tyr Phe Arg Arg Val Ala Ala Ala His Gln
250                 255                 260                 265

CTT CAC CAC TCG GAT AAA TTC CAC GGT GTT CCA TAT GGG CTC TTT CTC   929
Leu His His Ser Asp Lys Phe His Gly Val Pro Tyr Gly Leu Phe Leu
                270                 275                 280

GGA CCT AAG GAG CTT GAA GAA GTG GGG GGA CTA GAA GAA TTG GAG AAG   977
Gly Pro Lys Glu Leu Glu Glu Val Gly Gly Leu Glu Glu Leu Glu Lys
            285                 290                 295

GAG ATC AGT AAG AGA ATC AAA TCA TAC AAC AGG GTT CCA AAA          1019
Glu Ile Ser Lys Arg Ile Lys Ser Tyr Asn Arg Val Pro Lys
        300                 305                 310

TAATCAATTT AATGGGAGGA CCAATTTTTG GATCAATTTG TCAGTGTACA GAAACAATAG 1079

TGTTATTAAT GAAAAAAATA AATTATGAAT GCTTATGGGT GGATTACTGT TGTAAAGTTT 1139

ATGATGTTAA ATAATATAT                                             1158
```

SEQ ID NO: 2

SEQUENCE LENGTH: 311

SEQUENCE TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION

```
Met Ala Val Gly Leu Leu Ala Ala Ile Val Pro Lys Pro Phe Cys Leu
 1               5                  10                  15
```

```
Leu Thr Thr Lys Leu Gln Pro Ser Ser Leu Leu Thr Thr Lys Pro Ala
            20                  25                  30
Pro Leu Phe Ala Pro Leu Gly Thr His His Gly Phe Phe Asn Gly Lys
        35                  40                  45
Asn Arg Arg Lys Leu Asn Ser Phe Thr Val Cys Phe Val Leu Glu Glu
    50                  55                  60
Lys Lys Gln Ser Thr Gln Ile Glu Thr Phe Thr Asp Glu Glu Glu Glu
65                  70                  75                  80
Glu Ser Gly Thr Gln Ile Ser Thr Ala Ala Arg Val Ala Glu Lys Leu
                85                  90                  95
Ala Arg Lys Arg Ser Glu Arg Phe Thr Tyr Leu Val Ala Ala Val Met
            100                 105                 110
Ser Ser Phe Gly Ile Thr Ser Met Ala Val Met Ala Val Tyr Tyr Arg
        115                 120                 125
Phe Trp Trp Gln Met Glu Gly Gly Glu Val Pro Leu Ala Glu Met Phe
    130                 135                 140
Gly Thr Phe Ala Leu Ser Val Gly Ala Ala Val Gly Met Glu Phe Trp
145                 150                 155                 160
Ala Arg Trp Ala His Lys Ala Leu Trp His Ala Ser Leu Trp His Met
                165                 170                 175
His Glu Ser His His Arg Pro Arg Glu Gly Pro Phe Glu Leu Asn Asp
            180                 185                 190
Val Phe Ala Ile Ile Asn Ala Val Pro Ala Ile Ala Leu Leu Ser Phe
        195                 200                 205
Gly Phe Phe His Lys Gly Leu Val Pro Gly Leu Cys Phe Gly Ala Gly
    210                 215                 220
Leu Gly Ile Thr Val Phe Gly Met Ala Tyr Met Phe Val His Asp Gly
225                 230                 235                 240
Leu Val His Lys Arg Phe Pro Val Gly Pro Ile Ala Asp Val Pro Tyr
```

245 250 255

Phe Arg Arg Val Ala Ala Ala His Gln Leu His His Ser Asp Lys Phe

260 265 270

His Gly Val Pro Tyr Gly Leu Phe Leu Gly Pro Lys Glu Leu Glu Glu

275 280 285

Val Gly Gly Leu Glu Glu Leu Glu Lys Glu Ile Ser Lys Arg Ile Lys

290 295 300

Ser Tyr Asn Arg Val Pro Lys

305 310

SEQ ID NO: 3

SEQUENCE LENGTH: 8

SEQUENCE TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION

Phe Glu Leu Asn Asp Val Phe Ala

1 5

SEQ ID NO: 4

SEQUENCE LENGTH: 6

SEQUENCE TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION

His Asp Gly Leu Val His

1 5

SEQ ID NO: 5

SEQUENCE LENGTH: 18

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid (synthetic DNA)

SEQUENCE DESCRIPTION

TTYGARCTAA AYGAYGTN 18

SEQ ID NO: 6

SEQUENCE LENGTH: 17

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid (synthetic DNA)

SEQUENCE DESCRIPTION

CACGAYGGTC TNGTNCA 17

SEQ ID NO: 7

SEQUENCE LENGTH: 45

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid (synthetic DNA)

SEQUENCE DESCRIPTION

AACTGGAAGA ATTCGCGGCC GCAGGAATTT TTTTTTTTT TTTTT 45

SEQ ID NO: 8

SEQUENCE LENGTH: 21

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid (synthetic DNA)

SEQUENCE DESCRIPTION

TGGAAGAATT CGCGGCCGCA G 21

SEQUENCE LISTING

<110> Director General of National Institute of Fruit Tree Science,
Ministry of Agriculture, Forestry and Fisheries;
Bio-Oriented Technology Research Advancement Institution

<120> β-CAROTENE HYDROXYLASE GENE

<130> N.75714 GCW

<140> EP98309859.1
<141> 1998-12-02

<150> JP97/331936
<151> 1997-12-02

<160> 8

<170> PatentIn Ver. 2.0

<210> 1
<211> 1158
<212> DNA
<213> Citrus unshiu

<220>
<221> CDS
<222> (87)...(1019)

<400> 1

```
ccacaatcca cttcacatca actcttcctc ttttcaagtg cttttactct aaaacccaaa 60

acctcgtaaa caaacaaaac cccacc atg gcg gtc gga cta ttg gcc gcc ata  113
                               Met Ala Val Gly Leu Leu Ala Ala Ile
                                 1               5

gtc ccg aag ccc ttc tgt ctc ctc aca aca aaa ctt caa ccc tct tcg   161
Val Pro Lys Pro Phe Cys Leu Leu Thr Thr Lys Leu Gln Pro Ser Ser
 10                  15                  20                  25

ctc ctc aca aca aaa ccc gct ccc ctt ttt gcc cct ctc ggt acc cac   209
Leu Leu Thr Thr Lys Pro Ala Pro Leu Phe Ala Pro Leu Gly Thr His
                 30                  35                  40

cat ggc ttc ttt aat ggc aaa aac cga aga aaa ctc aac tct ttc acc   257
His Gly Phe Phe Asn Gly Lys Asn Arg Arg Lys Leu Asn Ser Phe Thr
             45                  50                  55

gta tgt ttt gtt tta gag gag aaa aaa caa agc acc cag atc gag act   305
Val Cys Phe Val Leu Glu Glu Lys Lys Gln Ser Thr Gln Ile Glu Thr
         60                  65                  70
```

```
ttc acg gac gag gag gag gag gag tcg ggt acc cag atc tcg act gct   353
Phe Thr Asp Glu Glu Glu Glu Glu Ser Gly Thr Gln Ile Ser Thr Ala
 75                  80                  85

gcc cgc gtg gcc gag aaa ttg gcg aga aag aga tcc gag agg ttc act   401
Ala Arg Val Ala Glu Lys Leu Ala Arg Lys Arg Ser Glu Arg Phe Thr
 90                  95                 100                 105

tat ctc gtt gct gcc gtc atg tct agt ttt ggt atc act tcc atg gct   449
Tyr Leu Val Ala Ala Val Met Ser Ser Phe Gly Ile Thr Ser Met Ala
                110                 115                 120

gtc atg gct gtt tat tac agg ttc tgg tgg caa atg gag ggt gga gag   497
Val Met Ala Val Tyr Tyr Arg Phe Trp Trp Gln Met Glu Gly Gly Glu
            125                 130                 135

gtg cct tta gct gaa atg ttt ggc aca ttt gct ctc tct gtt ggt gct   545
Val Pro Leu Ala Glu Met Phe Gly Thr Phe Ala Leu Ser Val Gly Ala
        140                 145                 150

gct gtg ggc atg gag ttt tgg gca cga tgg gct cat aaa gct ctg tgg   593
Ala Val Gly Met Glu Phe Trp Ala Arg Trp Ala His Lys Ala Leu Trp
    155                 160                 165

cat gct tct tta tgg cat atg cac gag tct cac cat cga cca aga gag   641
His Ala Ser Leu Trp His Met His Glu Ser His His Arg Pro Arg Glu
170                 175                 180                 185

ggt cct ttt gag cta aac gat gtg ttt gcc ata atc aac gca gtt cca   689
Gly Pro Phe Glu Leu Asn Asp Val Phe Ala Ile Ile Asn Ala Val Pro
                190                 195                 200

gcc ata gcc ctt ctc tct ttt ggc ttc ttc cac aaa ggc ctt gta cct   737
Ala Ile Ala Leu Leu Ser Phe Gly Phe Phe His Lys Gly Leu Val Pro
            205                 210                 215

ggt ctc tgc ttt ggt gct gga ctt ggc att acg gtg ttt ggg atg gcc   785
Gly Leu Cys Phe Gly Ala Gly Leu Gly Ile Thr Val Phe Gly Met Ala
        220                 225                 230

tac atg ttc gtc cac gat ggt ctc gtt cac aaa agg ttc cct gtg ggt   833
Tyr Met Phe Val His Asp Gly Leu Val His Lys Arg Phe Pro Val Gly
    235                 240                 245

ccc att gcc gac gtg cct tat ttc cgg aga gtc gct gcg gct cac cag   881
Pro Ile Ala Asp Val Pro Tyr Phe Arg Arg Val Ala Ala Ala His Gln
250                 255                 260                 265

ctt cac cac tcg gat aaa ttc cac ggt gtt cca tat ggg ctc ttt ctc   929
Leu His His Ser Asp Lys Phe His Gly Val Pro Tyr Gly Leu Phe Leu
                270                 275                 280

gga cct aag gag ctt gaa gaa gtg ggg gga cta gaa gaa ttg gag aag   977
Gly Pro Lys Glu Leu Glu Glu Val Gly Gly Leu Glu Glu Leu Glu Lys
            285                 290                 295
```

gag atc agt aag aga atc aaa tca tac aac agg gtt cca aaa 1019
Glu Ile Ser Lys Arg Ile Lys Ser Tyr Asn Arg Val Pro Lys
300 305 310

taatcaattt aatgggagga ccaatttttg gatcaatttg tcagtgtaca gaaacaatag 1079

tgttattaat gaaaaaaata aattatgaat gcttatgggt ggattactgt tgtaaagttt 1139

atgatgttaa ataatatat 1158

<210> 2
<211> 311
<212> PRT
<213> Citrus unshiu

<400> 2
Met Ala Val Gly Leu Leu Ala Ala Ile Val Pro Lys Pro Phe Cys Leu
1 5 10 15

Leu Thr Thr Lys Leu Gln Pro Ser Ser Leu Leu Thr Thr Lys Pro Ala
20 25 30

Pro Leu Phe Ala Pro Leu Gly Thr His His Gly Phe Phe Asn Gly Lys
35 40 45

Asn Arg Arg Lys Leu Asn Ser Phe Thr Val Cys Phe Val Leu Glu Glu
50 55 60

Lys Lys Gln Ser Thr Gln Ile Glu Thr Phe Thr Asp Glu Glu Glu Glu
65 70 75 80

Glu Ser Gly Thr Gln Ile Ser Thr Ala Ala Arg Val Ala Glu Lys Leu
85 90 95

Ala Arg Lys Arg Ser Glu Arg Phe Thr Tyr Leu Val Ala Ala Val Met
100 105 110

Ser Ser Phe Gly Ile Thr Ser Met Ala Val Met Ala Val Tyr Tyr Arg
115 120 125

Phe Trp Trp Gln Met Glu Gly Gly Glu Val Pro Leu Ala Glu Met Phe
130 135 140

Gly Thr Phe Ala Leu Ser Val Gly Ala Ala Val Gly Met Glu Phe Trp
145 150 155 160

Ala Arg Trp Ala His Lys Ala Leu Trp His Ala Ser Leu Trp His Met
165 170 175

His Glu Ser His His Arg Pro Arg Glu Gly Pro Phe Glu Leu Asn Asp
180 185 190

Val Phe Ala Ile Ile Asn Ala Val Pro Ala Ile Ala Leu Leu Ser Phe
195 200 205

```
Gly Phe Phe His Lys Gly Leu Val Pro Gly Leu Cys Phe Gly Ala Gly
    210                 215                 220

Leu Gly Ile Thr Val Phe Gly Met Ala Tyr Met Phe Val His Asp Gly
225                 230                 235                 240

Leu Val His Lys Arg Phe Pro Val Gly Pro Ile Ala Asp Val Pro Tyr
                245                 250                 255

Phe Arg Arg Val Ala Ala Ala His Gln Leu His His Ser Asp Lys Phe
            260                 265                 270

His Gly Val Pro Tyr Gly Leu Phe Leu Gly Pro Lys Glu Leu Glu Glu
        275                 280                 285

Val Gly Gly Leu Glu Glu Leu Glu Lys Glu Ile Ser Lys Arg Ile Lys
    290                 295                 300

Ser Tyr Asn Arg Val Pro Lys
305                 310
```

<210> 3
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Peptide designed based on the highly conserved amino acid sequence of β-carotene hydroxylase

<400> 3
```
Phe Glu Leu Asn Asp Val Phe Ala
  1               5
```

<210> 4
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Peptide designed based on the highly conserved amino acid sequence of β-carotene hydroxylase

<400> 4
```
His Asp Gly Leu Val His
  1               5
```

<210> 5
<211> 18
<212> DNA
<213> Artificial Sequence

<220>

<223> Synthetic DNA

<400> 5
ttygarctaa aygaygtn 18

<210> 6
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic DNA

<400> 6
cacgayggtc tngtnca 17

<210> 7
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic DNA

<400> 7
aactggaaga attcgcggcc gcaggaattt tttttttttt ttttt 45

<210> 8
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic DNA

<400> 8
tggaagaatt cgcggccgca g 21

Claims

1. A polypeptide which has β-carotene hydroxylase activity and which comprises an amino acid sequence selected from:

   (i) the amino acid sequence of SEQ ID NO:2;
   (ii) an amino acid sequence which is at least 70% homologous to the amino acid sequence of SEQ ID NO:2; and
   (iii)a fragment of sequence (i) or (ii) which is not a said sequence (ii).

2. A polypeptide according to claim 1 which is a recombinant protein selected from:

(a) a protein consisting of the amino acid sequence of SEQ ID NO:2; and
(b) a protein which consists of the amino acid sequence of SEQ ID NO:2 having deletion, substitution or addition of one or several amino acids.

3. A polynucleotide coding for a polypeptide as defined in claim 1 or 2.

4. A polynucleotide according to claim 3 which is a DNA.

5. A polynucleotide according to claim 4, which comprises the coding portion of the nucleotide sequence of SEQ ID NO:1.

6. A recombinant vector comprising a polynucleotide as defined in any one of claims 3 to 5.

7. A transformant which is transformed with a vector as defined in claim 6.

8. A method for preparing β-carotene hydroxylase and/or β-cryptoxanthin, which method comprises maintaining a transformant as defined in claim 7 under conditions such that the desired β-carotene hydroxylase and/or β-cryptoxanthin is expressed and recovering the expressed β-carotene hydroxylase and/or β-cryptoxanthin.

9. A method for preparing a β-carotene hydroxylase according to claim 8, comprising culturing the transformant in a medium and recovering the β-carotene hydroxylase from the resultant culture.

10. A method for preparing β-cryptoxanthin according to claim 8, comprising culturing the transformant in a medium and recovering β-cryptoxanthin from the resultant culture.

FIG. 1

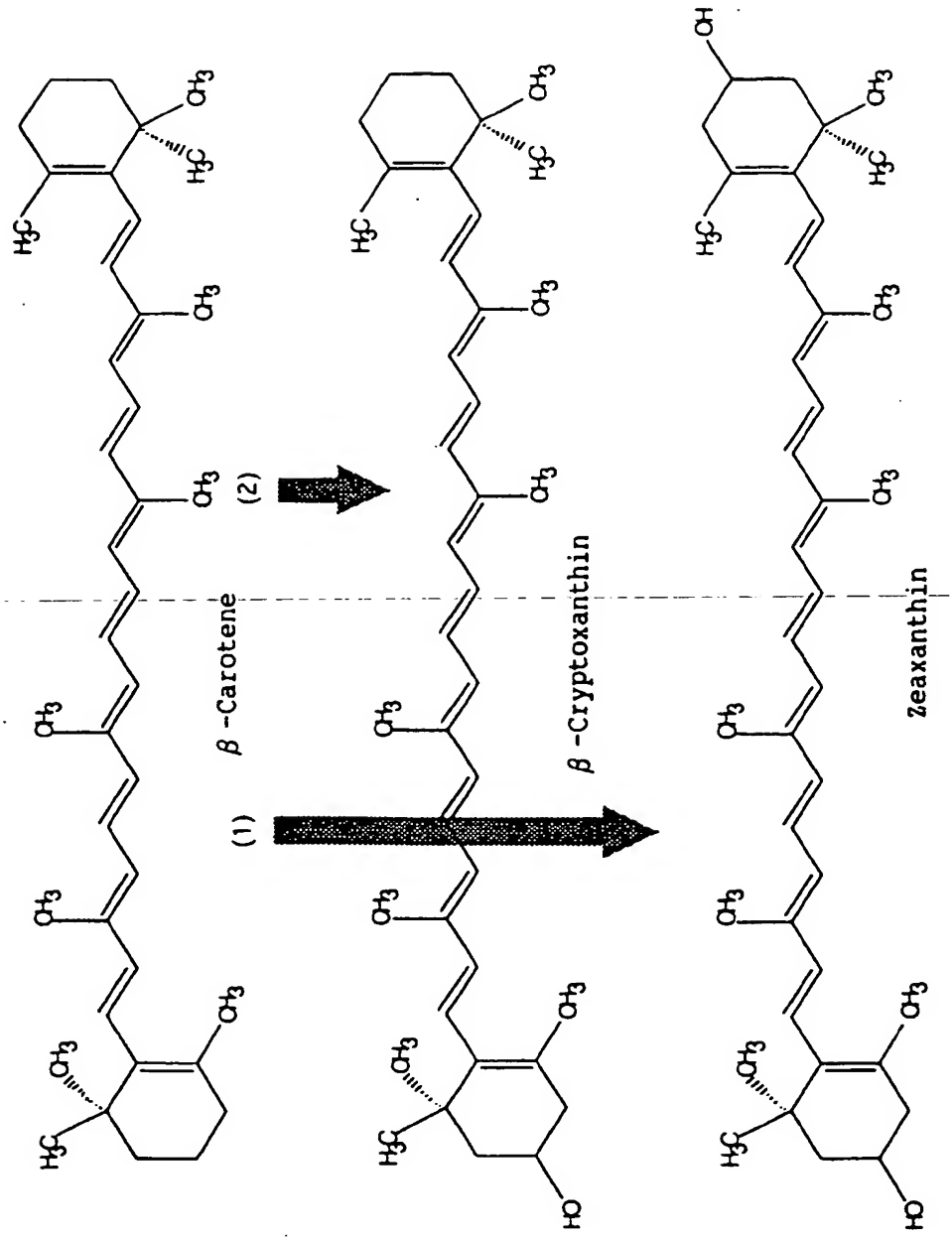


FIG. 2

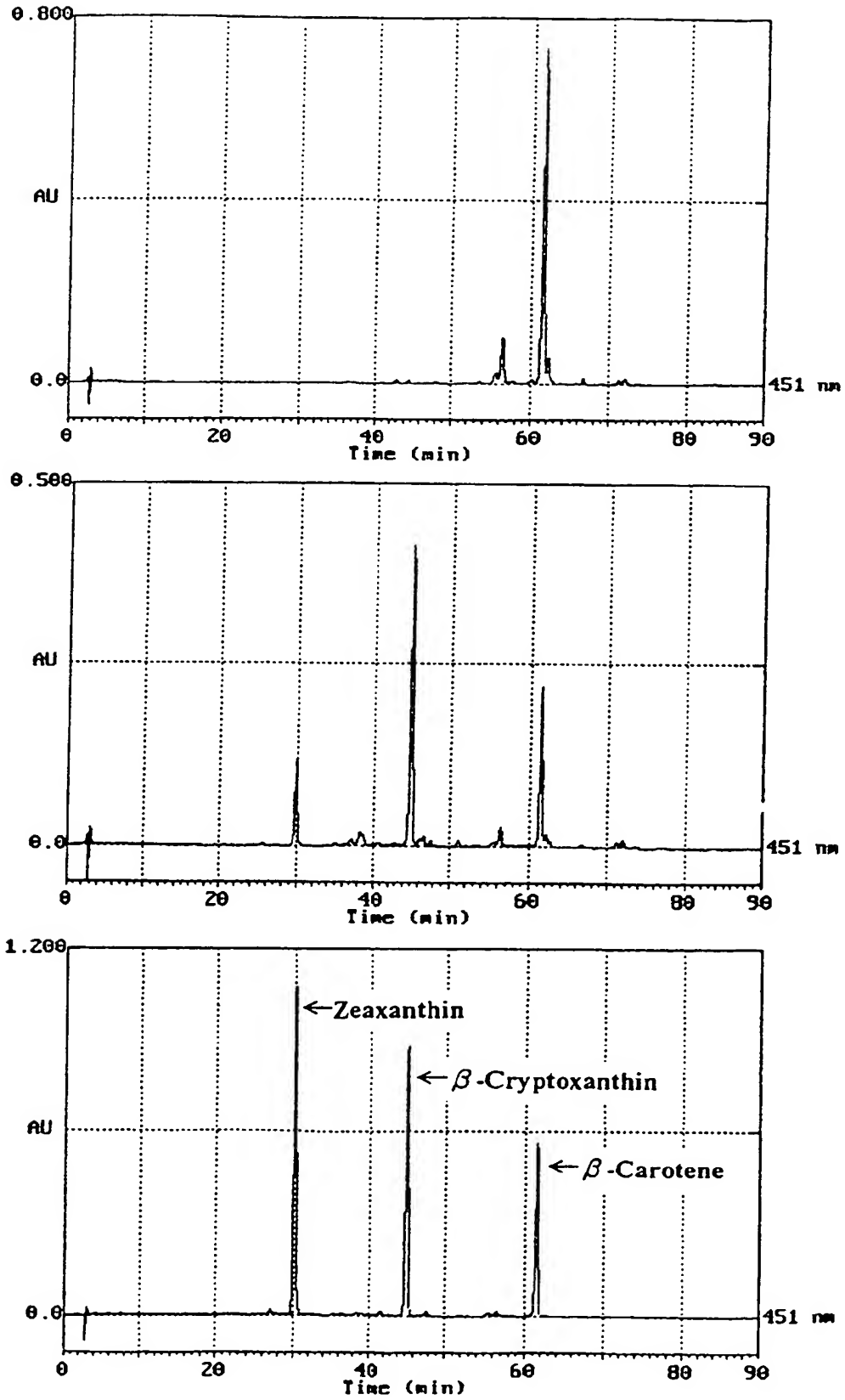


FIG. 3

```
Peptide
Score Table: Unitary Matrix
GAP Penalty: -4

CitBECH1                          1:MAVGLLAAIVPKPFCLLTTKLQPSSLLTTKPAPLFAPLGTHHGFFNGKNRRKLNSFTVCFVLEEKKQSTQIETFTDEEEEESGTQISTA-
Arabidopsis                       1:-----------------------.-FSSSSTDFRLR.PKSLSG.-SPSL.-FKR.S..Y.V..RR.NSP..NDERP.STS.TNA.DAEY
Agrobacterium aurantiacum-crtZ    1:----------------------------------------------------------------------------------------
Alicalgenes sp-crtZ               1:----------------------------------------------------------------------------------------
Erwinia herbicola crtZ            1:----------------------------------------------------------------------------------------
Erwinia uredovora-crtZ            1:----------------------------------------------------------------------------------------

CitBECH1                         91:-A-RVAEKLARKRSERFTYLVAAVMSSFGITSMAVMAVYYRFWWQMEGGEVPLAEMFGTFALSVGAAVGMEFWARWAHKALWHASLWHMH
Arabidopsis                      91:L.L.L....E..K...S...I..ML.....................S.......ISML.....................R........N..
Agrobacterium aurantiacum-crtZ   91:-----------------------------------------------------.TNFLIVVATVL.MELTAYSVHRWIMHGPLG.GW.
Alicalgenes sp-crtZ              91:-----------------------------------------------------.TQFLIVVATVL.MELTAYSVHRWIMHGPLG.GW.
Erwinia herbicola crtZ           91:-----------------------------------------------------MLVNSLIVILSVIAMEGIA.FTHRYIMHG-WG.RW.
Erwinia uredovora-crtZ           91:-----------------------------------------------------MLWIWNALIVF.TVIGMEVIA.LAHKYIMHG-WG.GW.
                                                                                                                  •  •

CitBECH1                        181:ESHHRPREGPFELNDVFAIINAVPAIALLSFGFFHKGLVPGLCFGAGLGITVFGMAYMFVHDGLVHKRFPVGPIADVPYFRRVAAAHQLH
Arabidopsis                     181:....K.............V..G...G...Y...N.....................I.....................L.K........
Agrobacterium aurantiacum-crtZ  181:K...EEHDHAL.K..LYGLVF..I.TV.FTV.WIWAPVLWW----IA..M..Y.LI.FVL.....Q.W.FRY.PRKG.A..LYQ..R..
Alicalgenes sp-crtZ             181:K...EEHDHAL.K..LYGVVF..L.TI.FTV.AYWWPVLWW----IA..M..Y.LI.FIL.....Q.W.FRY.PRRG....LYQ..R..
Erwinia herbicola crtZ          181:....T..K.V.....L..VVF.GV....IAV.TAGVWPLQW----I.C.M..Y.LL.FL.....Q.W.FHW.PRRG.LK.LYV..R..
Erwinia uredovora-crtZ          181:L...E..K.A..V..LY.VVF.ALS.L.IYL.STGMWPLQW----I.A.M.AY.LL.FM.....Q.W.FRY.PRKG.LK.LYM..RM.
                                     •••       • ••      •     •   •                • •  •  •   •••••• • •   •    •      ••  •

CitBECH1                        271:HSDKFHGVPYGLFLGPKELEEVGGLEELEKEISKRIKSYNRVPK-------
Arabidopsis                     271:.T...N.................N...D....R.....KKASGSGSSSSS
Agrobacterium aurantiacum-crtZ  271:.AVEGRDHCVSFGFIYAPPVDKLKQDLKMSGVLRAEAQERT----------------------------------------------
Alicalgenes sp-crtZ             271:.AVEGRDHCVSFGFIYAPPVDKLKQDLKRSGVLRPQDERPS-----------------------------------------------
Erwinia herbicola crtZ          271:.AVRGREGCVSFGFIYARKPADLQAILR.RHGRPPKRDAAKDRPDAASPSSSSPE-----------------------------------
Erwinia uredovora-crtZ          271:.AVRGKEGCVSFGFLYAPPLSKLQATLR.RHGARAGAARDAQGGEDEPASGK---------------------------------------
                                     •
```

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 933 427 A3

(12) EUROPEAN PATENT APPLICATION

(88) Date of publication A3:
19.01.2000 Bulletin 2000/03

(43) Date of publication A2:
04.08.1999 Bulletin 1999/31

(21) Application number: **98309859.1**

(22) Date of filing: **02.12.1998**

(51) Int. Cl.[7]: **C12N 15/53**, C12N 9/02, C12P 23/00

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **02.12.1997 JP 33193697**

(71) Applicants:
- **Director General of National Institute of Fruit tree Science, Ministry of Agriculture, Forestry and Fisheries**
  Tsukuba-shi, Ibaraki 305-0852 (JP)
- **Bio-oriented Technology Research Advancement Institution**
  Omiya-shi, Saitama 331-0044 (JP)

(72) Inventors:
- **Yano, Masamitsu**
  Shimizu-shi, Shizuoka 424-0205 (JP)
- **Omura, Mitsuo**
  Shizuoka-shi, Shizuoka 422-8021 (JP)
- **Ikoma, Yoshinori**
  Shimizu-shi, Shizuoka 424-0902 (JP)
- **Komatsu, Akira**
  Shimizu-shi, Shizuoka 424-0204 (JP)

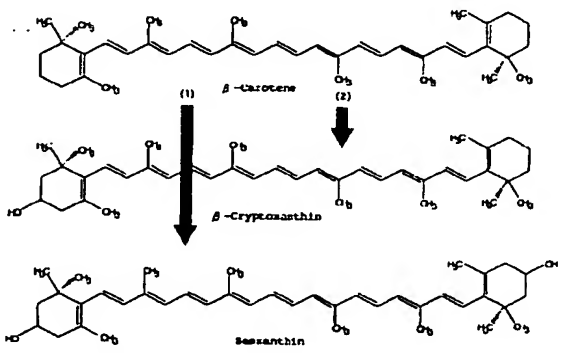
(74) Representative:
Woods, Geoffrey Corlett
J.A. KEMP & CO.
Gray's Inn
14 South Square
London WC1R 5LX (GB)

(54) Beta-carotene hydroxylase gene

(57) A polypeptide which has β-carotene hydroxylase activity comprises an amino acid sequence selected from:

(i) the amino acid sequence of SEQ ID NO:2;
(ii) an amino acid sequence which is at least 70% homologous to the amino acid sequence of SEQ ID NO:2; and
(iii)a fragment of sequence (i) or (ii) which is not a said sequence (ii).

FIG. 1



EP 0 933 427 A3

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 98 30 9859

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
X	WO 97 36998 A (UNIV MARYLAND) 9 October 1997 (1997-10-09) * SEQ ID NO 3 * * SEQ ID NO 4 * * page 13 - page 16 * * claims 15-23,27 * * page 12, paragraph 2 * * abstract * ---	1-4,6-8,10	C12N15/53 C12N9/02 C12P23/00
D,X	SUN Z ET AL: "Cloning and functional analysis of the beta-carotene hydroxylase of Arabidopsis thaliana." JOURNAL OF BIOLOGICAL CHEMISTRY, (1996 OCT 4) 271 (40) 24349-52. , XP002121880 * the whole document * ---	1,2,6-8,10 2-4,6-10	
Y	HUNDLE ET AL: "In vitro expression and activity of lycopene cyclase and beta-carotene hydroxylase from Erwinia herbicola" FEBS LETT, vol. 315, no. 3, 11 January 1993 (1993-01-11), pages 329-334, XP002121881 * figures 3-5 * * page 331 - page 334 * ---	2-4,6-10	TECHNICAL FIELDS SEARCHED (Int.Cl.6) C12N C12P
Y	RUTHER A ET AL: "Production of zeaxanthin in Escherichia coli transformed with different carotenogenic plasmids." APPLIED MICROBIOLOGY AND BIOTECHNOLOGY, (1997 AUG) 48 (2) 162-7. , XP002122719 * table 2 * * page 164-165 * ---	2-4,6-10	
A	WO 91 13078 A (AMOCO CORP) 5 September 1991 (1991-09-05) * examples 21-25 * --- -/--	2-8,10	

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
BERLIN	17 November 1999	ALCONADA RODRIG.., A

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 98 30 9859

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
A	EP 0 747 483 A (HOFFMANN LA ROCHE) 11 December 1996 (1996-12-11) * abstract * * examples 4,5 * * table 1 *	2-8,10	

TECHNICAL FIELDS SEARCHED (Int.Cl.6)

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
BERLIN	17 November 1999	ALCONADA RODRIG.., A

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 98 30 9859

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

17-11-1999

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9736998 A	09-10-1997	US 5744341 A	28-04-1998
		AU 1578497 A	22-10-1997
		CA 2250096 A	09-10-1997
		EP 0889952 A	13-01-1999
WO 9113078 A	05-09-1991	CA 2055447 A	03-09-1991
		EP 0471056 A	19-02-1992
		JP 5504686 T	22-07-1993
		US 5545816 A	13-08-1996
		US 5530188 A	25-06-1996
		US 5530189 A	25-06-1996
		US 5684238 A	04-11-1997
		US 5618988 A	08-04-1997
		US 5656472 A	12-08-1997
EP 0747483 A	11-12-1996	JP 9023888 A	28-01-1997

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82